

A medium-throughput SNP assay for detecting genetic variation in coding and non-coding portions of the red fox genome

Benjamin N. Sacks · Dag Inge Våge ·
Mark J. Statham

Received: 9 September 2009 / Accepted: 12 September 2009 / Published online: 24 September 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract SNPs are fast becoming the marker of choice in population genetics studies of model organisms. However, a lack of efficient means to assay sufficient numbers of SNPs in non-model organisms has prevented their widespread use for conservation-related applications. We established a SNP genotyping assay (including development of new SNP markers) for detection of genetic variation in coding and non-coding parts of the fox genome, using a 96-locus multiplex format. Coding SNPs were associated with Agouti and Mc1r coat-color alleles, phylogenetically basal mitochondrial mutations, sex markers, B-chromosome proto-oncogenes, and serotonin-receptor genes. Non-coding SNPs

($n = 51$) included 27 newly described microsatellite-linked SNPs. A panel of red foxes ($n = 96$) was used to characterize minor allele frequencies and Hardy–Weinberg equilibrium. Genotyping accuracy was 99% for 72 SNPs relative to sequenced and known-phenotype animals.

Keywords Coat-color · Red fox ·
Single nucleotide polymorphisms · SNPs · *Vulpes vulpes*

Electronic supplementary material The online version of this article (doi:10.1007/s12686-009-9106-8) contains supplementary material, which is available to authorized users.

B. N. Sacks (✉) · M. J. Statham
Canid Diversity and Conservation Laboratory, Center for
Veterinary Genetics, University of California, Davis, 1 Shields
Avenue/Old Davis Rd, Davis, CA 95616-8744, USA
e-mail: bnsacks@ucdavis.edu

B. N. Sacks
Department of Population Health and Reproduction,
University of California, Davis, 1 Shields Avenue/Old Davis Rd,
Davis, CA 95616-8744, USA

B. N. Sacks
Department of Biological Sciences, California State University,
Sacramento, 6000 J Street, Sacramento, CA 95819, USA

D. I. Våge
Department of Animal and Aquacultural Sciences, Norwegian
University of Life Sciences, 1432 Trondheim, Norway

D. I. Våge
Centre for Integrative Genetics (CIGENE), 1432 Trondheim,
Norway

Single nucleotide polymorphisms (SNPs) occur in coding and non-coding regions of the genome and are, in principle, the ideal marker for population genetics studies (Morin et al. 2004). However, because SNPs are typically biallelic, large numbers of markers are required for most applications, which has thus far prevented their widespread application in studies of non-model species, including most threatened and endangered species. While various approaches have been successfully used to discover moderate numbers of SNPs in non-model organisms (e.g., Sacks and Louie 2008), published genotyping assays used in conservation-related applications rarely involve more than a dozen or so multiplexed markers. As part of a broader research project aimed at conservation of threatened red fox (*Vulpes vulpes*) populations, we developed several new SNP markers and a multiplex assay for genotyping 96 coding and non-coding SNPs in red foxes. Here, we describe these markers along with the assay and its validation.

We used the Golden Gate platform and BeadXpress reader (Illumina, Inc., San Diego, California), which supported genotyping assays of up to 96 SNPs per well and a 96-well format (9,216 reactions; Fan et al. 2006). Sequences containing SNPs were submitted to Illumina's proprietary Assay Design Tool (<http://www.illumina.com>), and a subset of 96 loci was selected based on the manufacturer's design

scores (Figure S1). Reagents were synthesized by Illumina and transferred to the University of California at Davis, Genome Center, where genotyping reactions were carried out. Allele calls were made using BeadStudio software (Illumina, Inc.).

The assay was ultimately composed of 12 targeted “diagnostic” (coding or linked to coding) SNPs, 74 non-coding or randomly acquired SNPs (i.e., general genomic markers), and, because no additional polymorphic SNPs with sufficiently high design scores were available, 10 monomorphic “dummy” SNPs serving as negative controls. Diagnostic SNPs resided in coat-color and serotonin-receptor genes, mitochondrial DNA, and on sex chromosomes (Figure S1).

The three coat-color markers indicated mutant and wild-type alleles of the two genes known to determine the major coat-color morphs (e.g., red, cross, silver) in the red fox—Agouti gene and the melanocortin 1 receptor (Mc1R) gene (Våge et al. 1997, 2003). The two Agouti markers were designed to be used in concert to specify presence or absence of a large deletion associated with the standard silver phenotype. The “Agouti-silver” SNP was designed to amplify (but is monomorphic) when the deletion is present (in homozygous and heterozygous states) but not when the deletion is absent (homozygous wild type). Correspondingly, “Agouti-wild type” was designed to amplify (also monomorphic) when the wild-type allele is present (in homozygous and heterozygous states) but not when the deletion is present in both gene copies. The Mc1R-C125R marker detects a point mutation (SNP) in the exon of the Mc1R gene associated with the Alaska silver phenotype.

The mitochondrial cytochrome *b* SNPs were designed to differentiate major red fox clades as previously described in a worldwide survey (Aubry et al. 2009). The “Cytb-Holarctic” SNP distinguishes the two major clades (Holarctic, Nearctic) which split ~400 kya, while the “Cytb-East” marker distinguishes the Nearctic Eastern subclade from all others and, therefore, from the Nearctic Western subclades (split dating ~20 kya).

We used three SNPs from the serotonin 5-HT1A and 5-HT1B receptor genes (Kukekova et al. 2004). These SNPs were chosen as potential indicators for selection for tameness/habituation or aggression. They were discovered in noncoding regions of these genes as differing between two lines of red fox, one bred for tameness and the other randomly bred and, comparably more aggressive (Kukekova et al. 2004).

Two loci were used to assay for sex, one corresponding to the SRY gene on the non-recombining portion of the Y-chromosome (Genbank accession no. EU371736) and a second, the zinc finger protein genes Zfx and Zfy. Previously published SNPs differing between these homologues in various canids (Ortega et al. 2004) could not be used

because of an abundance of polymorphisms in the flanking regions. Therefore, we lined up publicly available sequences for the Zfx and Zfy genes of two distantly related canids, *Vulpes macrotis* and *Canis latrans* (Genbank accession nos. AY310319–AY310320; AY145847–848), and selected a SNP with base differing between chromosomal homologues but conserved between species.

The 74 genomic SNPs were selected from a suite of 123 SNPs, including 60 previously published SNPs based on screening of 9 individuals (Sacks and Louie 2008) and 63 additional microsatellite-linked SNPs presented here for the first time, discovered using a similar procedure except that we used a screening set of 4 red foxes ($n = 1$ each: United Kingdom, Idaho, California Sacramento Valley, California non-native).

Samples were run on a 96-well plate format and included 7 individuals from a Norwegian fur farm, each with a different coat-color genotype (and phenotype), representing all but two of the possible 9 genotypic combinations (Våge et al. 1997); 10 individuals that were sequenced during the SNP discovery processes, and 70 individuals chosen from a broad range of geographic locations including two distinct California populations used for assessing Hardy–Weinberg equilibrium in Genepop (Raymond and Rousset 1995). Accuracy was assessed relative to direct sequencing and phenotype (Table 1).

Seventy-two loci amplified in most individuals, while 24 loci were deemed unreliable (failing to amplify in all or most samples) and were excluded from further analyses (Table 1). Eighty-five samples genotyped at ≥ 60 of the 72 loci (average no. loci = 68.2, 95%) were considered successful and used in analyses. Accuracy was estimated at 98.8% overall ($n = 1,346$ comparisons) but varied somewhat across loci (Table 1). Accuracy of diagnostic SNPs was 99%, with the only potential errors occurring in the sex markers. However, all 85 sex genotypes agreed between Zfx/Zfy and SRY markers (48 M, 37 F), suggesting that discrepancies could reflect phenotyping errors by trappers in the field rather than genotyping errors. Accuracy was not tested directly in the B-chromosome and serotonin-receptor gene markers. However, the two successful serotonin-receptor markers approached Hardy–Weinberg equilibrium in the two California populations. Interestingly, allele frequencies of these markers, which were linked to genes potentially associated with tameness/habituation or aggression (Kukekova et al. 2004), differed between the two California populations. Despite previously reported polymorphism (Graphodatsky et al. 2005), the C-KIT SNPs were monomorphic in our sample.

Accuracy of genomic SNPs was estimated at 94.7%, somewhat lower than the diagnostic SNPs (Table 1). Of 45 successful genomic SNPs, 8 (18%) deviated significantly from Hardy–Weinberg equilibrium (i.e., population average

Table 1 Performance of 96 diagnostic (linked to or in targeted coding regions) and general genomic SNPs, including expected heterozygosity (H_e) and heterozygote deficit (F_{IS}) of SNPs in two California red fox populations, San Joaquin (SJ) valley ($n = 20$) and Sacramento (Sac) valley ($n = 24$)

SNP type	SNP name	H_e SJ valley	H_e Sac valley	F_{IS} (\bar{x})	% Correctly genotyped (n) ⁴	SNP or sequence reference
Diagnostic						
B-chromosome	C-KIT_73083 ¹	–	–	–	–	Graphodatsky et al. (2005)
B-chromosome	C-KIT_82570 ¹	–	–	–	–	Graphodatsky et al. (2005)
Coat-color	Agouti-Silver	–	–	–	100 (68)	This study; Våge et al. (2003)
Coat-color	Agouti-wild-type	–	–	–	100 (68)	This study; Våge et al. (2003)
Coat-color	Mc1R-C125R	–	–	–	100 (68)	Våge et al. (1997)
Mitochondrial	Cytb-East	–	–	–	100 (70)	This study; Aubry et al. (2009)
Mitochondrial	Cytb-Holarctic	–	–	–	100 (70)	This study; Aubry et al. (2009)
Serotonin rec	5-HT1A_1184CT	–	–	– ²	–	Kukekova et al. (2004)
Serotonin rec	5-HT1B_1684GT	0.10	0.49	0.02	–	Kukekova et al. (2004)
Serotonin rec	5-HT1B_174CT	0.50	0.12	–0.05	–	Kukekova et al. (2004)
Sex	Vv-SRY	–	–	–	93 (29)	This study
Sex	ZFX2	–	–	–	93 (29)	This study
Genomic						
Genomic	SL-10	0.47	0.13	–0.05	100 (3)	Sacks and Louie (2008)
Genomic	SL-11a	0.36	0.42	0.07	80 (5)	Sacks and Louie (2008)
Genomic	SL-11c	0.26	0	0.62*	100 (5)	Sacks and Louie (2008)
Genomic	SL-12a	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-12c	0.36	0.23	0.74*	100 (5)	Sacks and Louie (2008)
Genomic	SL-12d	0.36	0.40	0.19	100 (5)	Sacks and Louie (2008)
Genomic	SL-18	0.50	0.37	0.18	100 (5)	Sacks and Louie (2008)
Genomic	SL-20a	0.33	0.49	0.31	100 (5)	Sacks and Louie (2008)
Genomic	SL-21a	0.19	0.48	0.52	50 (2)	Sacks and Louie (2008)
Genomic	SL-22	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-23	0.41	0.42	0.37	100 (5)	Sacks and Louie (2008)
Genomic	SL-25	0.52	0.44	–0.05	100 (2)	Sacks and Louie (2008)
Genomic	SL-29	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-35a	0.38	0.49	0.01	50 (6)	Sacks and Louie (2008)
Genomic	SL-35b	0	0	– ³	100 (6)	Sacks and Louie (2008)
Genomic	SL-38	0.26	0.12	0.63*	100 (5)	Sacks and Louie (2008)
Genomic	SL-39	0.42	0.16	–0.09	100 (6)	Sacks and Louie (2008)
Genomic	SL-40	0.36	0.31	0.04	100 (6)	Sacks and Louie (2008)
Genomic	SL-41b	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-42	0.43	0.36	0.02	50 (6)	Sacks and Louie (2008)
Genomic	SL-43	0.47	0.44	0.23	100 (5)	Sacks and Louie (2008)
Genomic	SL-44	0.30	0.34	0.21	100 (5)	Sacks and Louie (2008)
Genomic	SL-47	0.47	0.28	0.13	100 (5)	Sacks and Louie (2008)
Genomic	SL-48a	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-48b	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-48c	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-48d	–	–	– ²	–	Sacks and Louie (2008)
Genomic	SL-49a	0	0	– ³	100 (5)	Sacks and Louie (2008)
Genomic	SL-49b	0.22	0.2	0.12	100 (5)	Sacks and Louie (2008)
Genomic	SL-50b	0.33	0.51	1.00*	100 (5)	Sacks and Louie (2008)
Genomic	SL-50c	0.23	0.49	0.45	80 (5)	Sacks and Louie (2008)
Genomic	SL-6	0.22	0.45	0.43	100 (5)	Sacks and Louie (2008)

Table 1 continued

SNP type	SNP name	H_e SJ valley	H_e Sac valley	F_{IS} (\bar{x})	% Correctly genotyped (n) ⁴	SNP or sequence reference
Microsat-linked	O-C01.424A	0.32	0.06	0.14	100 (3)	This study
Microsat-linked	O-C04.140A	–	–	$_{-2}$	–	This study
Microsat-linked	O-C04.140B	0	0	$_{-3}$	100 (1)	This study
Microsat-linked	O-C08.618A	0.51	0.51	0.05	100 (3)	This study
Microsat-linked	O-C08.618B	0.33	0.34	0.05	100 (3)	This study
Microsat-linked	O-CPH11	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH16A	0	0.25	0.51*	100 (3)	This study
Microsat-linked	O-CPH16B	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH18A	0.48	0.35	0.29	100 (3)	This study
Microsat-linked	O-CPH18C	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH19	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH2	0	0.28	0.41*	100 (3)	This study
Microsat-linked	O-CPH3A	0.14	0	–0.05	100 (2)	This study
Microsat-linked	O-CPH6	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH8A	0.10	0	–0.03	50 (2)	This study
Microsat-linked	O-CPH8E	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH8H	0.44	0	1.00*	100 (3)	This study
Microsat-linked	O-CPH9A	–	–	$_{-2}$	–	This study
Microsat-linked	O-CPH9B	0.51	0.51	0.11	100 (3)	This study
Microsat-linked	O-FH2001	0.23	0.46	0.48	100 (2)	This study
Microsat-linked	O-FH2004A	0.14	0	–0.05	100 (3)	This study
Microsat-linked	O-FH2004C	–	–	$_{-2}$	–	This study
Microsat-linked	O-FH2088D	–	–	$_{-2}$	–	This study
Microsat-linked	O-FH2100A	–	–	$_{-2}$	–	This study
Microsat-linked	O-FH2100B	0.09	0.33	0.30*	100 (3)	This study
Microsat-linked	O-FH2457	0.29	0.10	–0.09	100 (3)	This study
Microsat-linked	O-INRA021C	0.14	0.25	–0.10	100 (3)	This study
Microsat-linked	O-INU005	0.51	0.22	0.33	100 (3)	This study
Microsat-linked	O-INU030	0.51	0.36	0.46	100 (3)	This study
Microsat-linked	O-INU055A	–	–	$_{-2}$	–	This study
Microsat-linked	O-INU055D	0	0	$_{-3}$	100 (3)	This study
Microsat-linked	O-INU055E	–	–	$_{-2}$	–	This study
Microsat-linked	O-INU055G	0	0	$_{-3}$	100 (3)	This study
Microsat-linked	O-REN105L03A	0.51	0.38	–0.23	100 (3)	This study
Microsat-linked	O-REN105L03B	0.49	0.31	0.02	100 (3)	This study
Microsat-linked	O-REN162C04	0.18	0.15	–0.06	100 (3)	This study
Microsat-linked	O-REN247M23A	0.44	0.42	0.01	100 (3)	This study
Microsat-linked	O-REN247M23C	0.46	0.41	0.12	100 (3)	This study
Microsat-linked	O-REN247M23D	0.29	0.31	0.61	–	This study
Microsat-linked	O-REN54P11A	0	0	$_{-3}$	100 (3)	This study
Microsat-linked	O-REN54P11C	–	–	$_{-2}$	–	This study
Microsat-linked	O-REN64E19	–	–	$_{-2}$	–	This study
	Dummy 1 (SL-19)	–	–	–	100 (88)	Sacks and Louie (2008)
	Dummy 10 (SL-6)	–	–	–	100 (85)	Sacks and Louie (2008)
	Dummy 2 (SL-24)	–	–	–	97 (63)	Sacks and Louie (2008)
	Dummy 3 (HT1A)	–	–	–	100 (87)	Kukekova et al. (2004)
	Dummy 4 (HT1A)	–	–	–	100 (86)	Kukekova et al. (2004)
	Dummy 5 (HT1A)	–	–	–	100 (87)	Kukekova et al. (2004)

Table 1 continued

SNP type	SNP name	H_e SJ valley	H_e Sac valley	F_{IS} (\bar{x})	% Correctly genotyped (n) ⁴	SNP or sequence reference
	Dummy 6 (HT1A)	–	–	– ²		Kukekova et al. (2004)
	Dummy 7 (HT1A)	–	–	–	100 (84)	Kukekova et al. (2004)
	Dummy 8 (HT1A)	–	–	–	100 (89)	Kukekova et al. (2004)
	Dummy 9 (SL-5)	–	–	–	100 (85)	Sacks and Louie (2008)

¹ C-KIT SNPs reported by Graphodatsky et al. (2005) were monomorphic among all foxes used in this study

² SNP marker failed to amplify in the assay

³ Three SNPs monomorphic in the two California populations (O-INU055D, O-INU055G, O-REN54P11A) had relatively high heterozygosity in eight red foxes from the United Kingdom; three other SNPs monomorphic in the California samples (SL-35b, SL49a, O-C04.140B) had the other allele present (but rare) in another North American population

⁴ The genotyping accuracy was assessed for sex loci (from both markers together) based on a subset of individuals for which phenotypic sex was known, for coat-color loci (all three markers together) based on seven fur-farm foxes representing all but two possible genotypes (Mc1R heterozygous with Agouti heterozygous and homozygous wild type) and 61 wild-caught red-phase individuals, and for Cytb (both markers together) and genomic loci relative to a subset of individuals that was sequenced directly

* Deviation from Hardy–Weinberg equilibrium is significant; average P value between populations < 0.05

$P < 0.05$). All eight cases reflected heterozygote deficits, potentially indicating allelic dropout. Low DNA concentration of our samples could have contributed to errors; in most cases DNA concentration was <50 ng/ul (manufacturer recommended concentration) and in some cases was <1 ng/ul (unpublished data).

The accuracy estimates reported here, although generally conservative, were nonetheless sufficiently high for most population genetic applications. The ability to simultaneously run a suite of 51 genomic markers in concert with coat-color, sex, and phylogenetically informative mitochondrial markers (among others) makes this a useful tool. To our knowledge, this is the first of its kind available for a nonmodel wildlife species. In the future, additional markers can be added (e.g., to replace unsuccessful ones) further improving the utility of the red fox assay. The recent sequencing of the dog genome will also facilitate access to red fox genome sequence and SNP discovery in the future. More generally, our findings demonstrate the possibility of developing medium-throughout SNP genotyping assays for ecological and conservation-related applications.

Acknowledgments We thank C. Nicolette, V. Rashbrook, T. Kun, and M. Moore for technical support, C. Soulsbury and L. Dalén for European red fox samples. This work was funded in part by the Goethe Bequest (Biological Survey grant), California State University Sacramento, the University of California, Davis (UCD) Genome Center (Core Facility Pilot Project), California Department of Fish and Game/UCD Center for Population Biology, and the UCD, Veterinary Genetics Laboratory.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which

permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aubry KA, Statham MJ, Sacks BN, Perrine JD, Wisely SM (2009) Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. *Mol Ecol* 18:2668–2686
- Fan JB, Gunderson KL, Bibikova M, Yeakley JM, Chen J et al (2006) Illumina universal bead arrays. *Methods Enzymol* 410:57–73
- Graphodatsky AS, Kukekova AV, Yudkin DV et al (2005) The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosome Res* 13:113–122
- Kukekova AV, Trut LN, Oskina IN et al (2004) A marker set for construction of a genetic map of the silver fox (*Vulpes vulpes*). *J Hered* 95:185–194
- Morin PA, Luikart G, Wayne RK et al (2004) SNPs in ecology, evolution, and conservation. *Trends Ecol Evol* 19:208–215
- Ortega J, Franco MDR, Adams BA, Ralls K, Maldonado JE (2004) A reliable noninvasive method for sex-determination in the endangered San Joaquin kit fox (*Vulpes macrotis mutica*) and other canids. *Conserv Genet* 5:715–718
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
- Sacks BN, Louie S (2008) Using the dog genome to find SNPs in red foxes and other distantly related members of the Canidae. *Mol Ecol Resour* 8:35–49
- Våge DI, Lu DS, Klungland H, Lien S, Adalsteinsson S, Cone RD (1997) A non-epistatic interaction of Agouti and extension in the fox, *Vulpes vulpes*. *Nat Genet* 15:311–315
- Våge DI, Stavdal BE, Beheim J, Klungland H (2003) Why certain silver fox genotypes develop red hairs in their coat. *Scientific* 27:79–82